Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 2757–2760, April 1989
Genetics

Allele-specific enzymatic amplification of β-globin genomic DNA for diagnosis of sickle cell anemia

(genetic diseases/base-pair mismatch/DNA polymerase/oligodeoxyribonucleotide/polymerase chain reaction)

DAN Y. WU*, LUIS UGOZZOLI†, BIJAY K. PAL‡, AND R. BRUCE WALLACE*

*Department of Molecular Biochemistry, Beckman Research Institute of the City of Hope, Duarte, CA 91010; †Laboratorio di Immunogenetica, Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy; and ‡Department of Biological Sciences, California State Polytechnic University, Pomona, CA 91768

Communicated by Eugene Roberts, December 27, 1988 (received for review December 12, 1988)

ABSTRACT **A rapid nonradioactive approach to the diagnosis of sickle cell anemia is described based on an allele-specific polymerase chain reaction (ASPCR). This method allows direct detection of the normal or the sickle cell β-globin allele in genomic DNA without additional steps of probe hybridization, ligation, or restriction enzyme cleavage. Two allele-specific oligonucleotide primers, one specific for the sickle cell allele and one specific for the normal allele, together with another primer complementary to both alleles were used in the polymerase chain reaction with genomic DNA templates. The allele-specific primers differed from each other in their terminal 3′ nucleotide. Under the proper annealing temperature and polymerase chain reaction conditions, these primers only directed amplification on their complementary allele. In a single blind study of DNA samples from 12 individuals, this method correctly and unambiguously allowed for the determination of the genotypes with no false negatives or positives. If ASPCR is able to discriminate all allelic variation (both transition and transversion mutations), this method has the potential to be a powerful approach for genetic disease diagnosis, carrier screening, HLA typing, human gene mapping, forensics, and paternity testing.**

Sickle cell anemia is the prototype of a genetic disease caused by a single base-pair mutation, an A → T transversion in the sequence encoding codon 6 of the human β-globin gene. In homozygous sickle cell anemia, the substitution of a single amino acid (Glu → Val) in the β-globin subunit of hemoglobin results in a reduced solubility of the deoxyhemoglobin molecule and erythrocytes assume irregular shapes. The sickled erythrocytes become trapped in the microcirculation and cause damage to multiple organs.

Kan and Dozy (1) were the first to describe the diagnosis of sickle cell anemia in the DNA of affected individuals based on the linkage of the sickle cell allele to an *Hpa* I restriction fragment length polymorphism. Later, it was shown that the mutation itself affected the cleavage site of both *Dde* I and *Mst* II and could be detected directly by restriction enzyme cleavage (2, 3). Conner *et al.* (4) described a more general approach to the direct detection of single nucleotide variation by the use of allele-specific oligonucleotide hybridization. In this method, a short synthetic oligonucleotide probe specific for one allele only hybridizes to that allele and not to others under appropriate conditions.

All of the above approaches are technically challenging, require a reasonably large amount of DNA, and are not very rapid. The polymerase chain reaction (PCR) developed by Saiki *et al.* (5) provided a method to rapidly amplify small amounts of a particular target DNA. The amplified DNA could then be readily analyzed for the presence of DNA sequence variation (e.g., the sickle cell mutation) by allele-specific oligonucleotide hybridization (6), restriction enzyme cleavage (5, 7), ligation of oligonucleotide pairs (8, 9), or ligation amplification (10). PCR increased the speed of analysis and reduced the amount of DNA required for it but did not change the method of analysis of DNA sequence variation. In this paper, we investigated whether PCR could be done in an allele-specific manner such that the presence or absence of an amplified fragment provides direct determination of genotype.

PCR utilizes two oligonucleotide primers that hybridize to opposing strands of DNA at positions spanning a sequence of interest. A DNA polymerase [either the Klenow fragment of *Escherichia coli* DNA polymerase I (5) or *Thermus aquaticus* DNA polymerase (11)] is used for sequential rounds of template-dependent synthesis of the DNA sequence. Prior to the initiation of each new round, the DNA is denatured and fresh enzyme is added in the case of the *E. coli* enzyme. In this manner, exponential amplification of the target sequences is achieved. We reasoned that if the 3′ nucleotide of one of the primers formed a mismatched base pair with the template due to the existence of single nucleotide variation, amplification would take place with reduced efficiency. Specific primers would then direct amplification only from their homologous allele. After multiple rounds of amplification, the formation of an amplified fragment would indicate the presence of the allele in the initial DNA.

MATERIALS AND METHODS

Oligonucleotide Synthesis. Oligonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer by the phosphoramidite method. They were purified by electrophoresis on a urea/polyacrylamide gel followed by high-performance liquid chromatography as described (12).

Source and Isolation of Human DNA. All genomic DNA samples with the exception of the β-thalassemia DNA were isolated from the peripheral blood of appropriate individuals. The β-globin genotype of these individuals was previously determined by hybridization with allele-specific oligonucleotide probes (4) as well as by hemoglobin electrophoresis. Thalassemia major DNA was obtained from an Epstein–Barr virus-transformed lymphocyte cell line obtained from the National Institute of General Medical Sciences Human Genetic Mutant Cell Repository (Camden, NJ). Thalassemia DNA was isolated from the cultured cells. All DNA preparations were performed according to a modified Triton X-100 procedure followed by proteinase K and RNase A treatment (13). The average yield of genomic DNA was ≈25 μg per ml of blood.

PCR. Hβ14A (5′-CACCTGACTCCTGA) and BGP2 (5′-AATAGACCAATAGGCAGAG) at a concentration of 0.12 μM were used as the primer set for the amplification of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PCR, polymerase chain reaction; ASPCR, allele-specific PCR.

normal β-globin gene (*a* primer set). Similarly, 0.12 μM Hβ14S (5′-CACCTGACTCCTGT) and 0.12 μM BGP2 were used as the primer set for the amplification of the sickle cell gene (*s* primer set). Both primer sets directed the amplification of a 203-base-pair (bp) β-globin allele-specific fragment. As an internal positive control, all reaction mixtures contained an additional primer set for the human growth hormone gene comprised of 0.2 μM GHPCR1 (5′- TTCCCAACCATTCCCTTA) and 0.2 μM GHPCR2 (5′-GGATTTCTGTTGTGTTTC) (*hGH* primer set). GHPCR1 and GHPCR2 direct the amplification of a 422-bp fragment of the human growth hormone gene. All reactions were performed in a vol of 50 μl containing 50 mM KCl, 10 mM Tris·HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% (wt/vol) gelatin, template DNA (0.5 μg/ml), and 0.1 mM each dATP, dCTP, dGTP, and TTP. Reactions were carried out for 25 cycles at an annealing temperature of 55°C for 2 min, a polymerization temperature of 72°C for 3 min, and a heat-denaturation temperature of 94°C for 1 min on a Perkin-Elmer Cetus DNA thermal cycler. At the end of the 25 rounds, the samples were held at 4°C in the thermal cycler until removed for analysis.

Analysis of the PCR Products. An aliquot (15 μl) from each of the completed PCR reactions was mixed with 5 μl of 5× Ficoll loading buffer (1× = 10 mM Tris·HCl, pH 7.5/1 mM EDTA/0.05% bromophenol blue/0.05% xylene cyanol/3% Ficoll) and subjected to electrophoresis in a 1.5% agarose gel. Electrophoresis was performed in 89 mM Tris·HCl/89 mM borate/2 mM EDTA buffer for 2 hr at 120 V. At the completion of electrophoresis, the gel was stained in ethidium bromide (1.0 μg/ml) for 15 min, destained in water for 10 min, and photographed by ultraviolet trans-illumination.

RESULTS

Experimental Design. The scheme describing allele-specific PCR (ASPCR) is shown in Fig. 1. Primer P1 is designed such that it is complementary to allele 1 but the 3′-terminal nucleotide forms a single base-pair mismatch with the DNA sequence of allele 2 (Fig. 1*B*, *). Under appropriate annealing temperature and PCR conditions, there is normal amplification of the P1–P3 fragment with DNA templates containing allele 1 (homo- or heterozygous), while there is little or no amplification from DNA templates containing allele 2. In a similar way, a primer (P2) could be designed that would allow the specific PCR amplification of allele 2 but not allele 1 DNA.

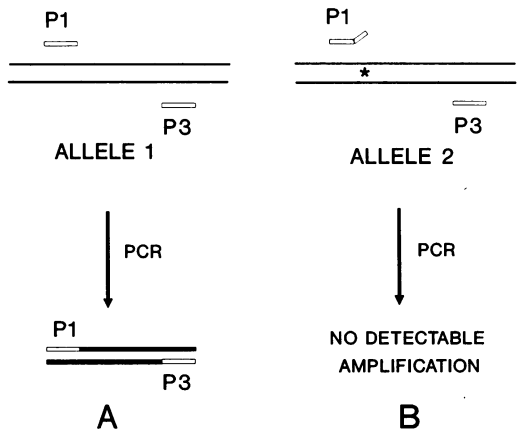

FIG. 1. Schematic representation of the ASPCR. P1 and P3, synthetic oligonucleotide primers that anneal to opposing strands of a single copy gene. P1 anneals to the region of a gene in the region of a DNA sequence variation such that its terminal 3′ nucleotide base pairs with the polymorphic nucleotide of the template. P1 is completely complementary to allele 1 (*A*) but forms a single base-pair mismatch with allele 2 at the 3′-terminal position due to one or more nucleotide differences relative to allele 1 (*B*).

We designed two 14-nucleotide-long allele-specific primers, Hβ14S and Hβ14A, complementary to the 5′ end of the sickle cell and normal β-globin genes, respectively. The oligonucleotide primers differ from each other by a single nucleotide at the 3′ end, Hβ14S having a 3′ T and Hβ14A having a 3′ A corresponding to the base pair affected by the sickle cell mutation. The oligonucleotide primer BGP2 (7) complementary to the opposite strand 3′ of the allele-specific primers was used as the second primer for PCR. The amplification product with these primer pairs was 203 bp. Also included in each reaction was a second pair of primers that directed the amplification of a 422-bp fragment of the human growth hormone gene. These primers were included as an internal positive control.

Discrimination Between the Normal and Sickle Cell Alleles. Genomic DNA was isolated from peripheral blood leukocytes of individuals of known β-globin genotypes (β^A/β^A, β^A/β^S, β^S/β^S). In addition, DNA was isolated from an Epstein–Barr virus-transformed cell line containing a homozygous deletion of the β-globin gene (β^{th}/β^{th}). DNA was subjected to 25 rounds of PCR using either the sickle cell-specific primer set (Hβ14S and BGP2) or the normal gene-specific primer set (Hβ14A and BGP2) using an annealing temperature of 55°C. The results are shown in Fig. 2*A*. It can be seen that a 203-bp fragment is observed using the sickle cell-specific primer set only with the β^A/β^S and β^S/β^S genomic DNA templates and not with the β^A/β^A genomic DNA templates. Conversely, the normal gene-specific primer set only gave rise to an amplification product with β^A/β^S and β^A/β^A genomic DNA templates. As expected, the thalassemia DNA did not give rise to a β-globin gene amplification product with either primer set. The internal growth hormone gene control gave rise to a 422-bp fragment in all samples, demonstrating that in no case was the absence of a globin-specific band due to a failure of the PCR.

In a single blind study, the DNA from 12 individuals with different β-globin genotypes was analyzed with the two primer sets. The results are shown in Fig. 2*B*. Individuals 1, 2, 3, and 5 are predicted to be β^A/β^A; individuals 6, 9, 10, and 11 are predicted to be β^S/β^S; and individuals 4, 7, 8, and 12 are predicted to be β^A/β^S. In each case, the genotype was correctly and unambiguously predicted from the pattern of fragment amplification (see legend to Fig. 2 for clinically diagnosed genotype).

DISCUSSION

The results presented above indicate the potential usefulness of ASPCR for sickle cell diagnosis. The method is rapid and the result is obtained without the use of radioactivity, since all that is required is to visualize the band on a gel with ethidium bromide staining. It should be possible to further improve the technique by elimination of the gel separation step. One strategy for this is shown in Fig. 3. As proposed recently by Yamane *et al.* (15), the two primers for the PCR could be labeled differently, one with biotin and one with a fluorescent group such as fluorescein or tetramethyl rhodamine. The product of the PCR could be captured on streptavidin-agarose and the presence of the amplified sequence could be detected with the fluorescence. In this case, if one allele-specific primer were labeled with one fluorescent group and the other were labeled with a different one, then the ASPCR could be done simultaneously.

In this study, we have used PCR primers that form either an A·A or a T·T mismatch. It is not clear that other mismatches will give equally effective discrimination. Since G·T mismatches are more stable than other mismatches (16), G·T should probably be avoided when designing primers.

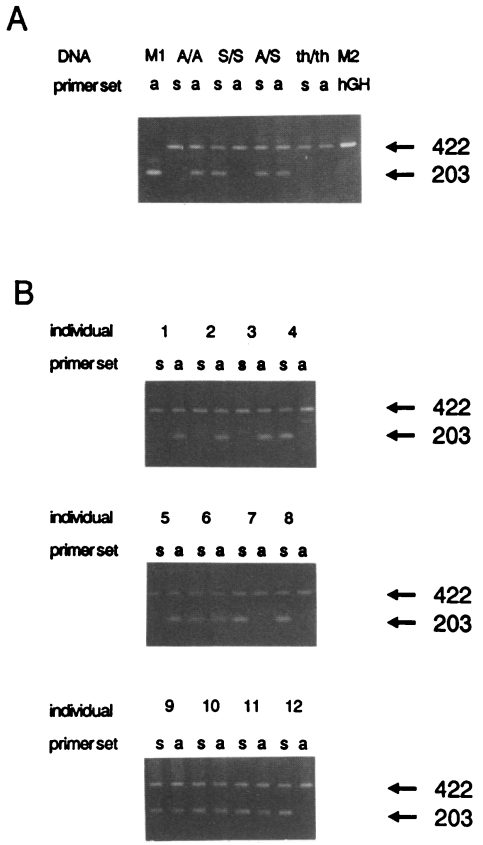

FIG. 2. (*A*) Identification of the normal (β^A) and the sickle cell (β^S) alleles by ASPCR. Normal (β^A/β^A), homozygous sickle cell (β^S/β^S), heterozygous sickle cell (β^A/β^S), and homozygous β-thalassemia (β^{th}/β^{th}) DNA samples (0.5 μg each) served as template using either the normal (*a* primer set) or the sickle cell (*s* primer set) for the ASPCRs. As an internal positive control, all reaction mixtures contained an additional primer set for the human growth hormone gene (*hGH* primer set) that directed the amplification of a 422-bp fragment of the human growth hormone gene. After amplification, 15 μl from each reaction mixture was subjected to electrophoresis in a 1.5% agarose gel for 2 hr at 120 V. Ethidium bromide staining of the agarose gel was used to detect PCR amplified fragments. Positive β-globin ASPCR can be identified by the presence of a 203-bp fragment using either the *a* or the *s* primer set reaction. As a marker for the globin-specific fragment, 0.3 μg of plasmid pHβ^A containing the normal human globin gene (β^A) was amplified with the *a* primer set alone (M1). As a marker for the growth hormone-specific fragment, 0.1 μg of plasmid pXGH5 containing a 3.8-kilobase fragment of the human growth hormone gene (14) was amplified with the growth hormone primer set (hGH) alone (M2). (*B*) A single blind trial using ASPCR to diagnose the β-globin genotype of genomic DNA samples. Genomic DNA samples from 12 individuals (4 each of normal, homozygous, and heterozygous sickle cell individuals) were randomly assigned numbers 1–12 by the hematology laboratory and blinded to the investigators. ASPCR was performed using both the normal (*a*) and the sickle cell-specific (*s*) primer sets as described above. Genotypes were identified as homozygous normal (β^A/β^A) if the single 203-bp fragment appears exclusively in the *a* primer set reaction, as homozygous sickle cell (β^S/β^S) if the 203-bp fragment appears only in the *s* primer set, or as heterozygous sickle cell trait (β^A/β^S) if the fragment appears in both reactions. The genotypes of these DNA samples were previously determined by hemoglobin electrophoresis (results not shown). The genotypes of the 12 individuals are as follows: 1, 2, 3, and 5, β^A/β^A; 6, 9, 10, and 11, β^A/β^S; 4, 7, 8, and 12, β^S/β^S.

This can be done by designing the primer so that it is complementary to the strand with which it forms an A·C mismatch. It may be possible to use a competition approach, as we have previously used to improve the discrimination provided by oligonucleotide hybridization probes (17). In this case, a competitive primer could be designed that was not able to prime, for example, by including in it a 3′ dideoxynucleotide or a 3′ ribonucleotide that has been oxidized. A mixture of a labeled allele-specific primer complementary to allele 1 plus an unlabeled priming-defective primer complementary to allele 2 should then allow the specific amplification of allele 1.

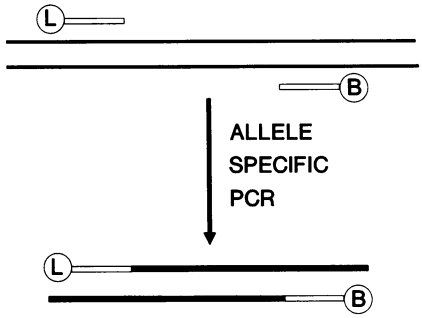

FIG. 3. Schematic representation of a dual labeling system suitable for the detection of the ASPCR products. One of the oligonucleotide primers is labeled at the 5′ end with a fluorescent group such as fluorescein or tetramethyl rhodamine (L) and the other primer is labeled with biotin (B). The ASPCR amplification product would therefore have the 5′ end labeled on both strands. The biotin is suitable for capturing the amplified fragment on a streptavidin-agarose column, while the fluorescent group is suitable for measuring the amount of fragment produced.

The ability of an oligonucleotide to prime on a DNA template is governed by two kinetic variables: the rate at which the annealed primer dissociates from the template before initiating polymerization (r_{off}) and the rate at which the DNA polymerase extends the primer (r_{pol}). Efficient priming in PCR should take place whenever $r_{pol} > r_{off}$, the addition of the first few nucleotides to the primer then greatly stabilizing the oligonucleotide–template complex and allowing continued extension of the primer. For a given primer r_{pol} is an intrinsic property of the polymerase. Studies with *E. coli* DNA polymerase I have suggested that this polymerase may be able to discriminate between primers that either do or do not form a mismatch with the template at the 3′-terminal nucleotide (18). In this case, r_{pol} for the mismatched primer was slower than r_{pol} for the perfectly matched primer. For the present study, we designed the allele-specific primers such that the allele-specific nucleotide in the template was complementary to the 3′-terminal nucleotide of the primer. In this way, the 3′ nucleotide of the primer specific for one allele would form a mismatch with the other allele. This design allows one to take advantage of the difference between r_{pol} of the perfectly matched and mismatched primers as well as to optimize primer concentration, priming temperature, primer length, and primer sequence, all of which will affect the difference in the r_{off} for the two allele-specific primers.

We reasoned that a set of conditions should exist such that $r_{pol} > r_{off}$ for the perfectly matched primer, while $r_{pol} < r_{off}$ for the mismatched primer. The results shown here clearly demonstrate this to be true. In our study, the allele-specific primers were 14 nucleotides long. We found (data not shown) that discrimination between the β^A and β^S alleles was not possible at low annealing temperatures (e.g., 44°C and 50°C). Presumably the short length of the oligonucleotides as well as the high annealing temperature combined to provide the discrimination.

Taq polymerase is well suited for using ASPCR for the discrimination of two alleles that differ by a single nucleotide because it lacks a 3′ → 5′ exonuclease activity (19). Such an activity would correct the mismatched base pair in the mismatched primer–template complex and then permit efficient priming with the one-nucleotide-shorter primer. Since

the specificity of the ASPCR is determined in the initial several cycles of PCR, the fact that the primer remains uncorrected enhances the discrimination of the reaction. PCR is an exponential reaction; the yield of product is very dependent on the efficiency of each round (5). Only very minor changes in the efficiency of each round of amplification have profound effects on the overall yield after many rounds. For example, if the efficiency of the reaction with the perfectly matched primer is 90% and with the mismatched primer is 60%, there would be 73-fold more product produced in the reaction with perfectly matched primer than with the mismatched primer.

The ASPCR should find application in the fields of genetic diagnosis, carrier screening, HLA typing, and any other nucleic acid-based diagnostic in which the precise DNA sequence of the priming site is diagnostic for the target. In the case of HLA typing, recent advances have used PCR amplification followed by allele-specific oligonucleotide hybridization for the determination of *DR*, *DQ*, and *DP* alleles (6, 20–22). It should be possible to use ASPCR for the direct analysis of HLA types.

We have recently proposed a process for the simultaneous determination of multiple polymorphic loci based on the concept of producing locus-specific amplification products each with a unique length (23). In such a system, since ASPCR would produce allele-specific products, the simultaneous analysis of the genotype of the target DNA at multiple loci should be possible.

This work was supported by Grant DCB-8515365 from the National Science Foundation (R.B.W.). D.Y.W. is a M.D./Ph.D. candidate at Loma Linda University. R.B.W. is a member of the Cancer Center of the City of Hope (NIH CA33572). L.U. is a fellow of AIRC (Associazione Italiana per la Ricerca sul Cancro).

1. Kan, Y. W. & Dozy, A. M. (1978) *Lancet* **ii**, 910–912.
2. Geever, R. F., Wilson, L. B., Nallaseth, F. S., Milner, P. F., Bittner, M. & Wilson, J. T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5081–5085.
3. Chang, J. C. & Kan, Y. W. (1982) *N. Engl. J. Med.* **307**, 30–32.
4. Conner, B. J., Reyes, A. A., Morin, C., Itakura, K., Teplitz, R. L. & Wallace, R. B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 278–282.
5. Saiki, R. K., Scharf, S., Falcona, F., Mullis, K., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
6. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1986) *Nature (London)* **324**, 163–166.
7. Chehab, F. F., Doherty, M., Cai, S., Kan, Y. W., Cooper, S. & Rubin, E. M. (1987) *Nature (London)* **329**, 293–294.
8. Landegren, U., Kaiser, R., Sanders, J. & Hood, L. (1988) *Science* **241**, 1077–1080.
9. Wu, D. Y. & Wallace, R. B. (1989) *Gene*, in press.
10. Wu, D. Y. & Wallace, R. B. (1989) *Genomics*, in press.
11. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
12. Dembek, P., Miyoshi, K. & Itakura, K. (1981) *J. Am. Chem. Soc.* **103**, 706–708.
13. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5759–5763.
14. Selden, R. F., Howie, K. B., Rowe, M. E., Goodman, H. M. & Moore, D. D. (1986) *Mol. Cell Biol.* **6**, 3173–3179.
15. Yamane, A., Nakagami, S., Kawasoe, T. & Miyoshi, K. (1988) *Nucleic Acids Res.* **20**, 91.
16. Kidd, V. J., Wallace, R. B., Itakura, K. & Woo, S. L. C. (1983) *Nature (London)* **304**, 230–234.
17. Nozari, G., Rahbar, S. & Wallace, R. B. (1986) *Gene* **43**, 23–28.
18. Atkinson, M. R., Deutscher, M. P., Kornberg, A., Russel, A. F. & Moffet, J. G. (1969) *Biochemistry* **8**, 4897–4904.
19. Tindall, K. R. & Kunkel, T. A. (1988) *Biochemistry* **27**, 6008–6013.
20. Morel, P. A., Dorman, J. S., Todd, J. A., McDevitt, H. O. & Trucco, M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8111–8115.
21. Angelini, G., Bugawan, T., Delfino, L., Erlich, H. & Ferrara, G. B. (1988) *Hum. Immunol.* **23**, 77.
22. Scharf, S., Saiki, R. & Erlich, H. (1988) *Hum. Immunol.* **23**, 143.
23. Skolnick, M. H. & Wallace, R. B. (1988) *Genomics* **2**, 273–279.